Sisteme sisteme multi-agent (proiect)

Student: Daniel-Cristian-Marian Țăpuși

Data: 18 mai 2024

Cuprins

[1. Principiul de modelare conform formalismului SMA a problemei dilema prizonierului – agenti componenti (tipuri standard si instantieri) & diagrama secventiala 1](#_Toc135587749)

[2. Modalitatea de luare a deciziei 1](#_Toc135587750)

[3. Realizarea protocolului de interactiune 1](#_Toc135587751)

[4. Configuratii de rulare 1](#_Toc135587752)

[5. Capturi de ecran din Sniffer 1](#_Toc135587753)

[6. Concluzii 1](#_Toc135587754)

# 1. Principiul de modelare conform formalismului SMA a problemei dilema prizonierului – agenti componenti (tipuri standard si instantieri) & diagrama secventiala

Agenti componenti:

1. Arbitru

Tip: Agent centralizat care controleaza fluxul jocului si monitorizeaza deciziile prizonierilor

Roluri:

* + Solicita deciziile de la prizonieri
  + Primeste deciziile
  + Proceseaza deciziile conform regulilor stabilite
  + Informeaza prizonierii despre rezultate

1. Prisoner1 și Prisoner2

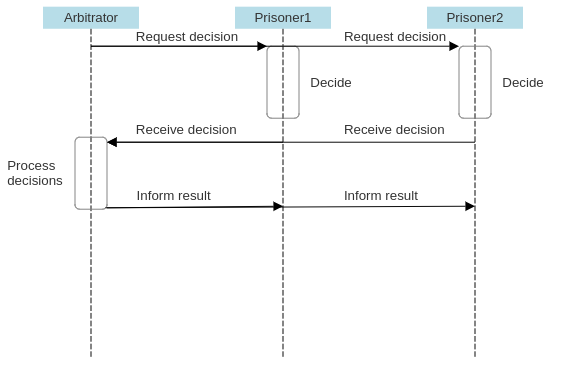
Tip: Agenti autonomi care iau decizii bazate pe strategia proprie

Roluri:

* + Primeste solicitari de la arbitrator
  + Decide sa coopereze sau sa tradeze
  + Trimite decizia inapoi la arbitrator

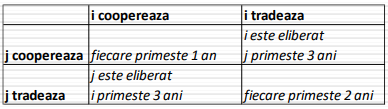
Instantieri:

Arbitrator este instantiat o singura data, gestionand intregul joc. Prisoner1 si Prisoner2 sunt instantiati individual, fiecare avand capacitatea de a lua decizii independent.



# 2. Modalitatea de luare a deciziei

Prizonierii sunt pusi in situatia de a alege intre a coopera (C) si a trada (T). Fiecare prizonier isi face alegerea fara a sti ce va decide celalalt prizonier. Deciziile sunt influentate de structura de plata a jocului, unde diferite combinatii de decizii rezulta in diferite recompense sau pedepse.



Strategii Comune:

* Tit-for-Tat: Un prizonier cooperează în prima rundă, apoi replică decizia adversarului din runda anterioară
* Always Defect: Prizonierul alege întotdeauna să trădeze, indiferent de comportamentul celuilalt prizonier
* Always Cooperate: Prizonierul alege întotdeauna să coopereze, indiferent de comportamentul celuilalt prizonier
* Random: Prizonierul decide la întâmplare între cooperare și trădare
* Grudger: Prizonierul începe prin a coopera și continuă să coopereze până când celălalt prizonier trădează, după aceea, prizonierul va continua să trădeze pentru restul jocului
* Pavlov: Cunoscută și sub numele de „Win-Stay, Lose-Shift”, această strategie implică repetarea deciziei anterioare dacă rezultatul a fost satisfăcător și schimbarea acesteia în caz contrar

Procesul de Decizie:

Fiecare prizonier evalueaza posibilele rezultate ale alegerilor de cooperare sau tradare, bazandu-se pe recompensele si pedepsele stabilite. Prizonierii pot incerca sa anticipeze decizia celuilalt bazandu-se pe strategia proprie sau pe informatiile disponibile despre comportamentul adversarului. Dupa evaluare, fiecare prizonier trimite decizia sa arbitrului, care va procesa ambele decizii pentru a determina rezultatul rundei.

Arbitrul nu influenteaza direct deciziile, dar este responsabil pentru a solicita deciziile, a le primi si a comunica rezultatele procesarii acestora, asigurand astfel ca jocul se desfasoara conform regulilor stabilite.

Prin urmare, modalitatea de luare a deciziei in Dilema Prizonierului implica evaluarea strategica a recompenselor si pedepselor, anticiparea deciziilor adversarului si adaptarea comportamentului bazat pe informatiile disponibile si pe structura de plata a jocului.

# 3. Realizarea protocolului de interactiune

Protocolul de interactiune in cadrul simularii Dilemei Prizonierului este esential pentru a asigura o interactiune coerenta si sistematica intre prizonieri. Acest protocol defineste regulile si mecanismele prin care agentii comunica, iau decizii si evalueaza rezultatele interactiunilor. Structura protocolului poate fi impartita in trei componente principale:

A. Initializarea Jocului

La inceputul simularii, fiecare prizonier este initializat cu o strategie specifica, si sunt stabilite regulile de baza ale interactiunii, cum ar fi numarul de runde si recompensele pentru diferitele combinatii de decizii (coopereaza/tradeaza). Acesta este momentul in care se configureaza si parametrii de mediu, iar prizonierii sunt informati despre regulile jocului.

B. Desfasurarea Rundelor

Protocolul de interactiune pentru fiecare runda de joc include urmatorii pasi:

1. Luarea Deciziilor: Fiecare prizonier, folosind metoda chooseNextDecision(), alege sa coopereze sau sa tradeze in functie de strategia adoptata. Decizia poate depinde de istoricul interactiunilor anterioare cu adversarul sau poate fi determinata aleatoriu, conform strategiei stabilite
2. Comunicarea Deciziilor: Decizicile sunt comunicate simultan pentru a evita influentarea deciziei unui prizonier de catre alegerea celuilalt. Acest lucru asigura ca niciun prizonier nu are un avantaj nejustificat prin cunoasterea deciziei adversarului inainte de a face propria alegere
3. Evaluarea Rezultatelor: Dupa ce ambele decizii sunt facute si comunicate, se calculeaza recompensele conform matricei de payoff a jocului. De exemplu, daca ambii prizonieri coopereaza, fiecare ar putea primi o recompensa; daca unul tradeaza si celalalt coopereaza, tradatorul primeste o recompensa mai mare, iar celalalt o penalizare
4. Actualizarea Scorurilor si a Statelor: Scorurile sunt actualizate in functie de rezultatele fiecarei runde

C. Concluzia Jocului

La sfarsitul jocului, dupa ce toate rundele au fost completate, scorurile finale sunt calculate si comparate pentru a determina strategia cea mai eficienta sau pentru a extrage invataminte despre comportamentul prizonierilor. Rezultatele pot fi folosite pentru analiza ulterioara sau pentru ajustarea strategiilor in jocuri viitoare.

# 4. Configuratii de rulare

1. AlwaysCooperate – TitforTat (50 - 50)

Rezultat: Daca un jucator coopereaza mereu, iar celalalt incepe cooperand si continua sa raspunda la fel, ambele parti vor coopera continuu, rezultand intr-un scor egal.

2. AlwaysDefect – Grudger (98 – 101)

Rezultat: Grudger va coopera in prima runda, iar AlwaysDefect va profita de aceasta, castigand un punct in plus fata de Grudger. Dupa aceea, amandoi vor continua sa tradeze, mentinand un scor relativ egal, dar cu un mic avantaj pentru AlwaysDefect datorita primei runde.

3. Pavlov – Random (110 – 20)

Rezultat: Pavlov este eficient in a "pedepsi" si "recompensa" comportamentul Random, adaptandu-se rapid si exploatand inconsistenta Random pentru a maximiza scorul.

4. Random – AlwaysDefect (131 - 38)

Rezultat: Random va avea unele runde de cooperare, care vor fi exploatate de AlwaysDefect. Tradarea constanta a lui AlwaysDefect va asigura ca acesta obtine puncte constant, in timp ce Random pierde puncte in runde de cooperare.

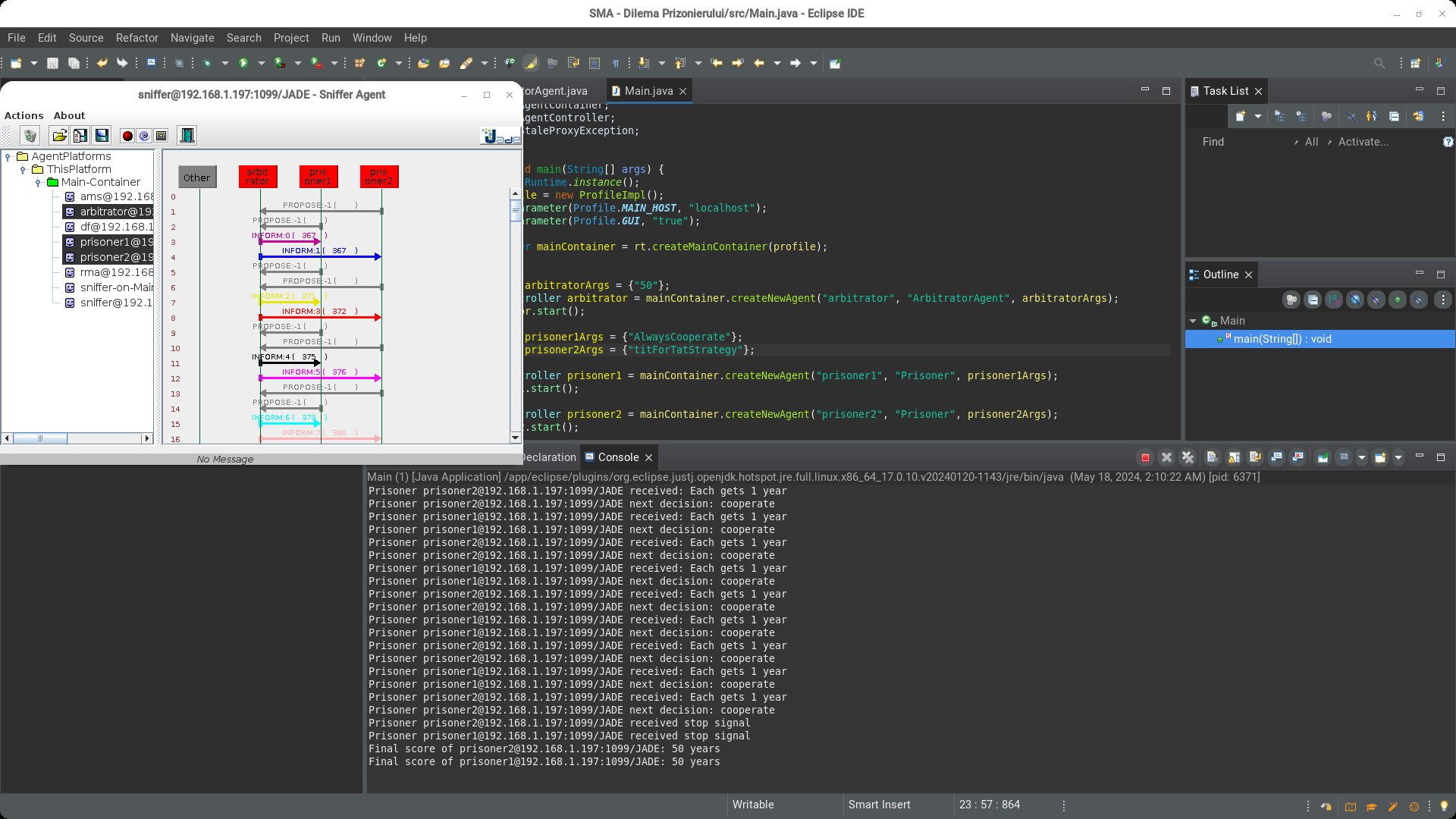
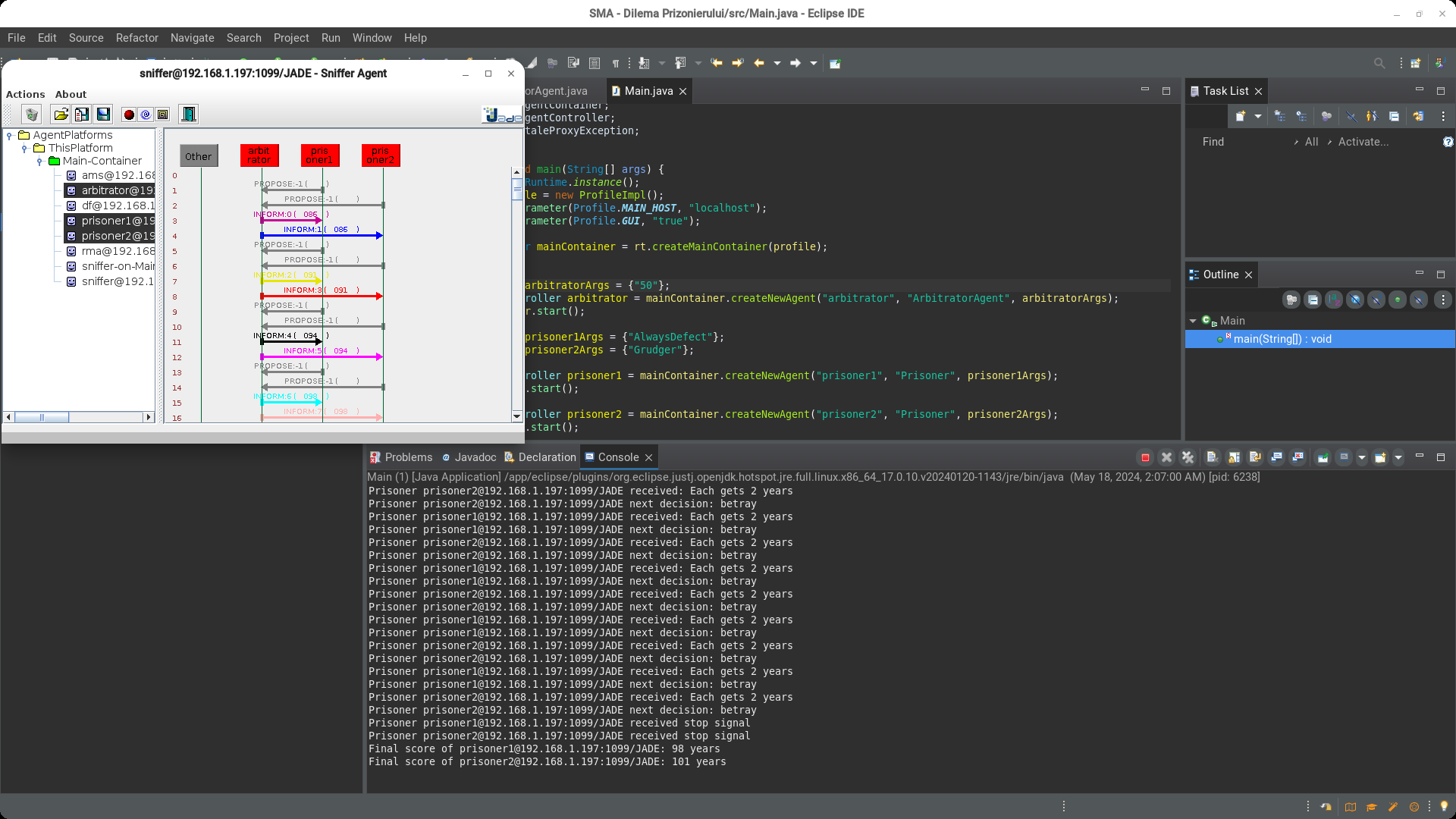
5. AlwaysDefect – TitforTat (101 - 98)

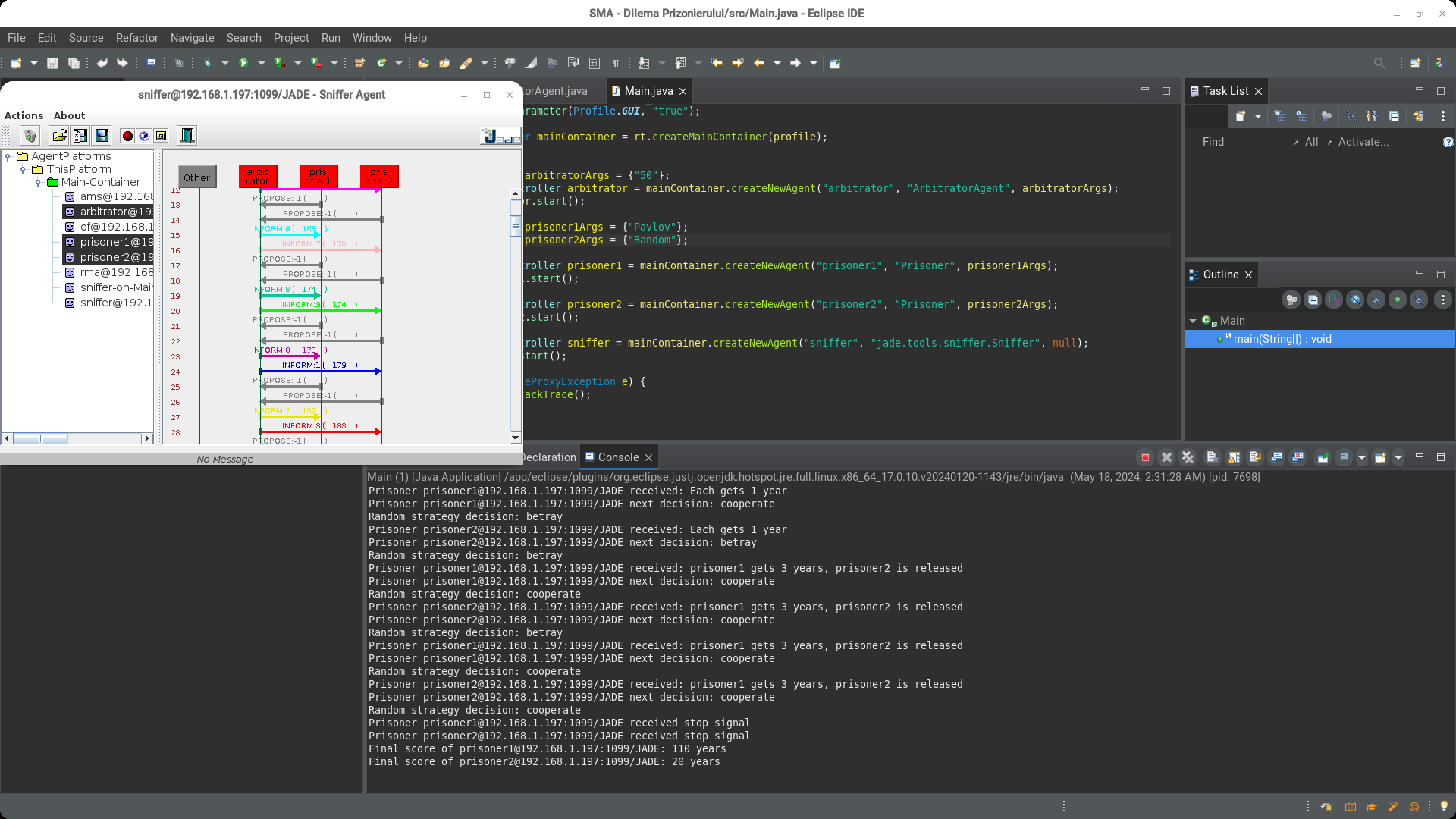
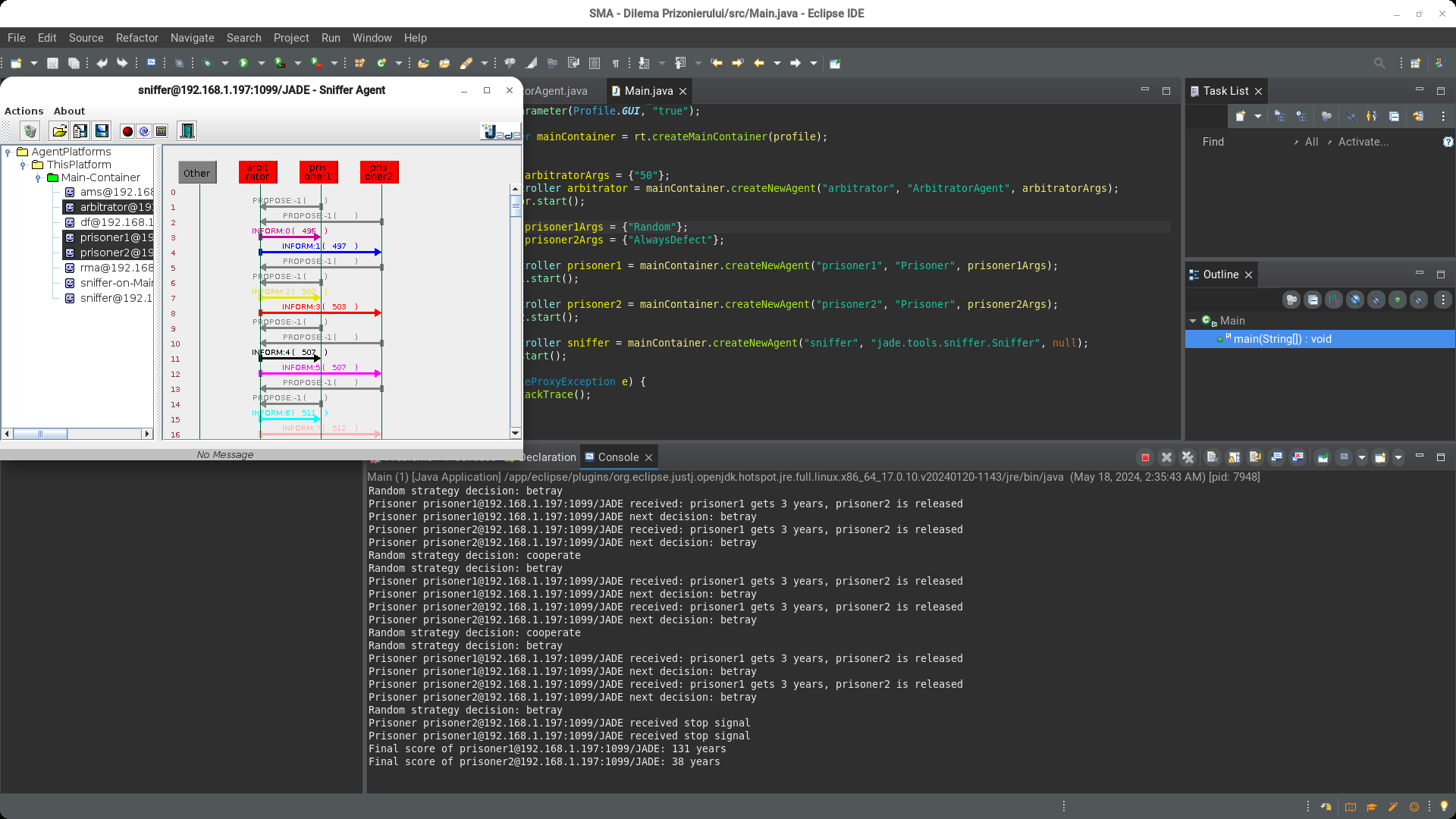
Rezultat: TitforTat va coopera in prima runda si va fi tradat de AlwaysDefect. De aici inainte, TitforTat va raspunde cu tradare, rezultand intr-un scor apropiat, dar AlwaysDefect va avea un avantaj minor datorita primei runde.

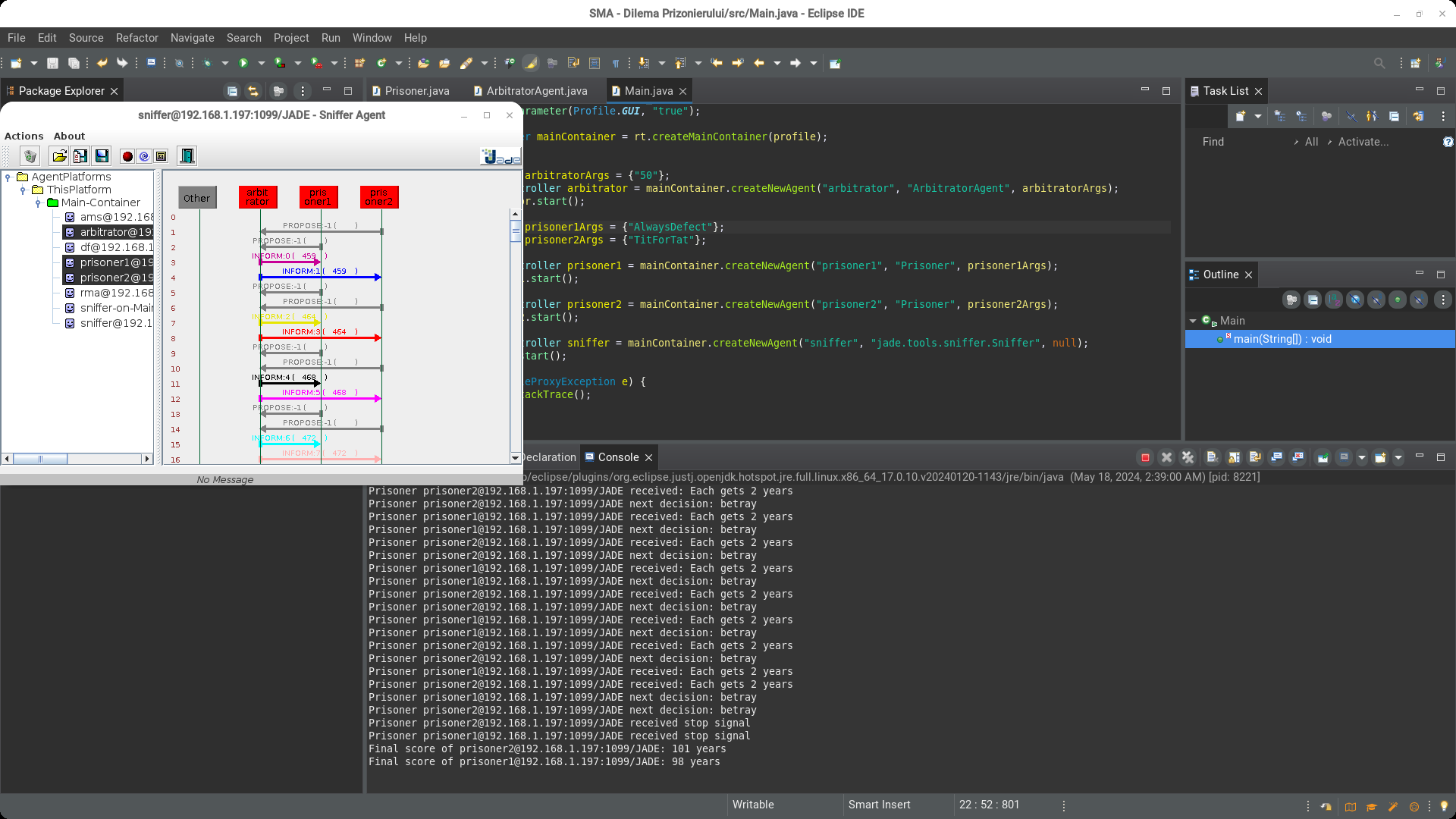
6. Random vs TitforTat (77 - 77)

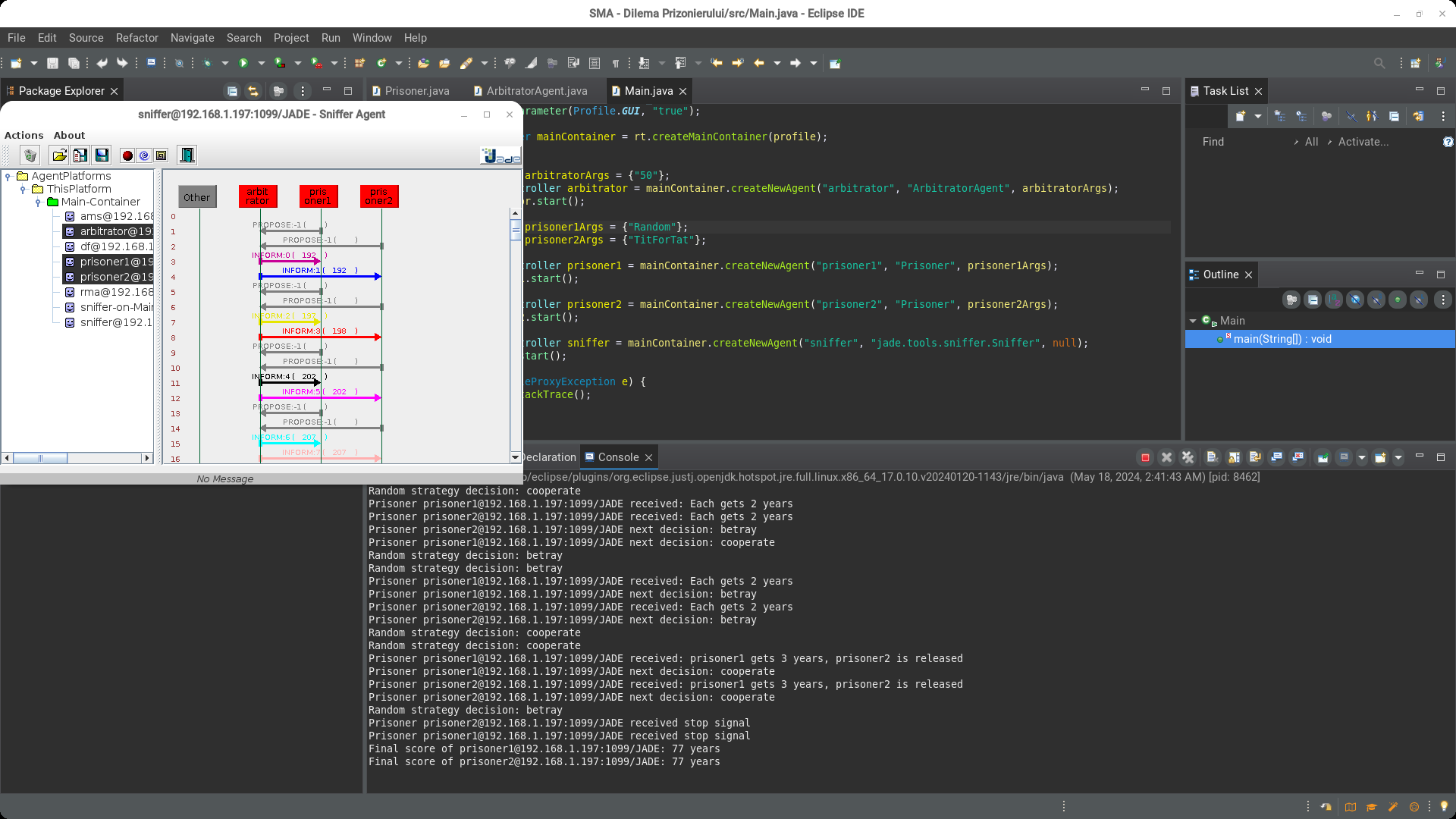
Rezultat: TitforTat va incepe prin a coopera si va replica comportamentul Random. Acest lucru poate duce la un scor egal, deoarece Random alterneaza intre cooperare si tradare, iar TitforTat raspunde in consecinta, echilibrand rezultatul final.

# 5. Capturi de ecran din Sniffer

AlwaysCooperate – TitforTat (50 - 50)AlwaysDefect – Grudger (98 – 101)

Pavlov – Random (110 – 20)Random – AlwaysDefect (131 - 38)

AlwaysDefect – TitforTat (101 - 98)

Random vs TitforTat (77 - 77)

# 6. Concluzii

Proiectul analizat a demonstrat clar superioritatea strategiei Tit-for-Tat (TFT) in contextul Dilemei Prizonierului repetata, subliniind eficienta sa in promovarea cooperarii si gestionarea conflictelor. Concluziile studiului pot fi rezumate astfel:

Eficienta Tit-for-Tat:

Promovarea cooperarii reciproce: Tit-for-Tat a inceput fiecare interactiune prin cooperare, stabilind un fundament pentru reciprocitate si incredere. Aceasta este cruciala pentru construirea unui mediu colaborativ, esential nu doar in teoria jocurilor, ci si in relatiile interpersonale si internationale.

Adaptabilitate:

TFT raspunde direct la actiunile adversarului. Daca adversarul coopereaza, TFT raspunde in kind, mentinand un mediu reciproc benefic. Daca adversarul tradeaza, TFT pedepseste aceasta tradare prin tradare reciproca in runda urmatoare, descurajand astfel comportamentele negative si limitand pierderile.

Simplu si previzibil:

Simplitatea strategiei TFT o face usor de inteles si de anticipat de catre adversari, ceea ce poate duce la o dinamica de joc mai stabila si la decizii mai informate din partea celorlalti jucatori.

Eficienta pe termen lung:

In comparatie cu alte strategii care pot avea succes pe termen scurt prin exploatarea adversarilor, TFT se dovedeste a fi eficienta pe termen lung, promovand un echilibru si reducand volatilitatea scorurilor intre jucatori.

Tit-for-Tat nu numai ca a demonstrat o eficienta remarcabila in cadrul acestui proiect, ci ofera si o perspectiva valoroasa asupra importantei adaptabilitatii si reciprocitatii in relatiile complexe. Prin promovarea cooperarii si prin raspunsul echilibrat la tradare, Tit-for-Tat se afirma ca o strategie superioara, recomandata pentru orice situatie in care cooperarea reciproca este esentiala pentru succesul pe termen lung.